PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, 15/53, 15/54, A01H 5/00, 5/10

A1

(11) International Publication Number: WO 96/01905

(43) International Publication Date: 25 January 1996 (25.01.96)

(21) International Application Number: PCT/US95/08501

(22) International Filing Date: 6 July 1995 (06.07.95)

(30) Priority Data:
08/272,917 8 July 1994 (08.07.94) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/272,917 (CIP)
Filed on 8 July 1994 (08.07.94)

(71) Applicant *(for all designated States except US)*: E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: FALCO, Saverio, Carl [US/US]; 1902 Millers Road, Arden, DE 19810-4032 (US).

(74) Agents: SIEGELL, Barbara, C. et al.; E.I. Du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

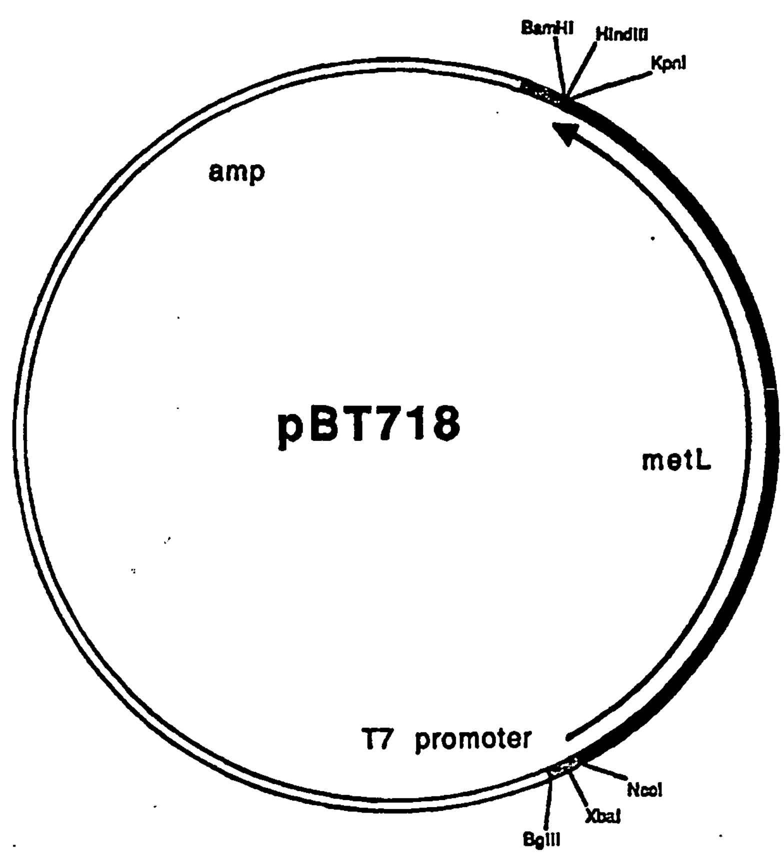
(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CHIMERIC GENES AND METHOD FOR INCREASING THE THREONINE CONTENT OF THE SEEDS OF PLANTS

(57) Abstract

A chimeric gene encoding a bifunctional feedback-insensitive aspartokinase-homoserine dehydrogenase (AK-HDH), is operably linked to a plant chloroplast transit sequence, and to plant seed-specific regulatory sequences. The chimeric gene is transformed into plants wherein increased levels of free threonine accumulate in the seeds.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SI	Slovenia
SK	Slovakia
SN	Senegal
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

TITLE

CHIMERIC GENES AND METHOD FOR INCREASING THE THREONINE CONTENT OF THE SEEDS OF PLANTS

TECHNICAL FIELD

This invention relates to a chimeric gene encoding a bifunctional feedback-insensitive aspartokinase-homoserine dehydrogenase (AK-HDH), which is operably linked to a plant chloroplast transit sequence, and to plant seed-specific regulatory sequences. A method for its use to produce increased levels of threonine in the seeds of transformed plants is provided.

BACKGROUND OF THE INVENTION

Human food and animal feed, derived from many grains, are deficient in essential amino acids, such as lysine, the sulfur amino acids methionine and cysteine, threonine and tryptophan. In corn, threonine is the third most limiting amino acid, after lysine and the sulfur amino acids, for the dietary requirements of many animals.

While considerable effort has been directed at increasing the lysine and sulfur amino acid content of crops, little has been done to attempt to increase the threonine content. Mutant corn and barley lines that had elevated whole-kernel threonine concentrations were isolated from cells grown in culture by selecting for growth in the presence of inhibitory concentrations of lysine plus threonine [Hibberd et al. (1982) *Proc. Natl. Acad. Sci. USA 79*:559-563, Bright et al. (1982) *Nature 299*:278-279]. However, agronomically-acceptable cultivars have never been derived from these lines and commercialized.

A preferable approach would be to increase the production and accumulation of the threonine in agronomically-acceptable cultivars, via genetic engineering technology. Threonine, along with methionine, lysine and isoleucine, are amino acids derived from aspartate. The first step in the biosynthetic pathway is the phosphorylation of aspartate by the enzyme aspartokinase (AK), and this enzyme has been found to be an important target for regulation of the pathway in many organisms. The aspartate family pathway is also believed to be regulated at the branch-point reactions. Thus, for threonine the reduction of aspartyl β-semialdehyde by homoserine dehydrogenase (HDH) may be an important point of control.

Galili and co-workers have reported on the introduction of an *E. coli* lysC gene that encodes a lysine-sensitive AK enzyme into tobacco cells via transformation [Galili et al. (1992) Eur. Patent Appl. 91119328.2; Shaul et al. (1992) *Plant Physiol. 100*:1157-1163]. Expression of the *E. coli* enzyme resulted in small increases in the levels of free threonine in the leaves and seeds of

transformed plants, but effects on the total threonine content were too small to be detected. Falco isolated a mutant of the *E. coli* lysC gene, which encoded a lysine-insensitive AK. Falco linked the gene to the bean phaseolin promoter and a plant chloroplast transit sequence, and transformed tobacco and canola with the chimeric gene. Expression of multiple copies of this chimeric gene in the seeds of transformed plants lead to a modest increase in the total threonine content of the seeds [World Patent Publication WO 93/19190]. Although the last reference provides a means to achieve a small threonine increase in seeds, there is a need for better chimeric genes and methods to increase the threonine content of seeds further.

SUMMARY OF THE INVENTION

This inventions concerns a chimeric gene wherein a nucleic acid fragment encoding a bi-functional protein with aspartokinase and homoserine dehydrogenase activities, both of which are substantially insensitive to end-product inhibition, is operably linked to a plant chloroplast transit sequence and to a seed-specific regulatory sequence. In a preferred embodiment, the nucleic acid fragment comprises the *E. coli* metL gene. This invention also concerns a plant comprising in its genome the chimeric gene described above and seeds obtained from that plant.

This invention further relates to a method for increasing the threonine content of the seeds of plants comprising:

(a) transforming plant cells with one of the chimeric genes described above;

(b) growing fertile mature plants from the transformed plant cells obtained from step (a) under conditions suitable to obtain seeds; and

(c) selecting from the progeny seed of step (b) for those seeds containing increased levels of threonine compared to untransformed seeds.

Also disclosed are seeds obtained by this method and plants obtained from such seeds.

BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE DESCRIPTIONS

The invention can be more fully understood from the following detailed description and the accompanying drawings and the sequence descriptions which form a part of this application.

Figure 1 shows a map of plasmid pBT718.

Figure 2 shows a map of plasmid pBT726.

Figure 3 shows a map of plasmid pBT727.

Figure 4 shows a map of plasmid pBT728.

Figure 5 shows a map of plasmid pBT733.

Figure 6 shows a map of plasmid pBT766.

SEQ ID NOS:1-2 were used in Example 1 as PCR primers to isolate and modify the *E. coli* metL gene.

SEQ ID NOS:3-8 were used in Example 2 to create a corn chloroplast transit sequence and link the sequence to the *E. coli* metL gene.

SEQ ID NO:9 was used in Example 4 to create a soybean chloroplast transit sequence and link the sequence to the *E. coli* metL gene.

DETAILED DESCRIPTION OF THE INVENTION

The teachings below describe nucleic acid fragments, chimeric genes and procedures useful for increasing the accumulation of threonine in the seeds of transformed plants, as compared to levels of threonine in untransformed plants.

In the context of this disclosure, a number of terms shall be utilized. As used herein, the term "nucleic acid" refers to a large molecule which can be single-stranded or double-stranded, composed of monomers (nucleotides) containing a sugar, phosphate and either a purine or pyrimidine. A "nucleic acid fragment" is a fraction of a given nucleic acid molecule. In higher plants, deoxyribonucleic acid (DNA) is the genetic material while ribonucleic acid (RNA) is involved in the transfer of the information in DNA into proteins. A "genome" is the entire body of genetic material contained in each cell of an organism. The term "nucleotide sequence" refers to a polymer of DNA or RNA which can be single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases capable of incorporation into DNA or RNA polymers.

As used herein, "essentially similar" refers to DNA sequences that may involve base changes that do not cause a change in the encoded amino acid, or which involve base changes which may alter one or more amino acids, but do not affect the functional properties of the protein encoded by the DNA sequence. It is therefore understood that the invention encompasses more than the specific exemplary sequences. Modifications to the sequence, such as deletions, insertions, or substitutions in the sequence which produce silent changes that do not substantially affect the functional properties of the resulting protein molecule are also contemplated. For example, alteration in the gene sequence which reflect the degeneracy of the genetic code, or which result in the production of a chemically equivalent amino acid at a given site, are contemplated; thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in

substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a biologically equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the protein molecule would also not be expected to alter the activity of the protein. In some cases, it may in fact be desirable to make mutants of the sequence in order to study the effect of alteration on the biological activity of the protein. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products. Moreover, the skilled artisan recognizes that "essentially similar" sequences encompassed by this invention are also defined by their ability to hybridize, under stringent conditions (0.1X SSC, 0.1% SDS, 65°C), with the sequences exemplified herein.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding) and following (3' non-coding) the coding region. "Native" gene refers to the gene as found in nature with its own regulatory sequences. "Chimeric" gene refers to a gene comprising heterogeneous regulatory and coding sequences. "Endogenous" gene refers to the native gene normally found in its natural location in the genome. A "foreign" gene refers to a gene not normally found in the host organism but that is introduced by gene transfer.

"Coding sequence" refers to a DNA sequence that codes for a specific protein and excludes the non-coding sequences.

"Initiation codon" and "termination codon" refer to a unit of three adjacent nucleotides in a coding sequence that specifies initiation and chain termination, respectively, of protein synthesis (mRNA translation). "Open reading frame" refers to the amino acid sequence encoded between translation initiation and termination codons of a coding sequence.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript. "Messenger RNA (mRNA) refers to RNA that can be translated into protein by the cell. "cDNA" refers to a double-stranded DNA, one strand of which is complementary to and derived from mRNA by reverse transcription. "Sense" RNA refers to RNA transcript that includes the mRNA.

As used herein, suitable "regulatory sequences" refer to nucleotide sequences located upstream (5'), within, and/or downstream (3') to a coding sequence, which control the transcription and/or expression of the coding sequences, potentially in conjunction with the protein biosynthetic apparatus of the cell. These regulatory sequences include promoters, translation leader sequences, transcription termination sequences, and polyadenylation sequences.

"Promoter" refers to a DNA sequence in a gene, usually upstream (5') to its coding sequence, which controls the expression of the coding sequence by providing the recognition for RNA polymerase and other factors required for proper transcription. A promoter may also contain DNA sequences that are involved in the binding of protein factors which control the effectiveness of transcription initiation in response to physiological or developmental conditions. It may also contain enhancer elements.

An "enhancer" is a DNA sequence which can stimulate promoter activity. It may be an innate element of the promoter or a heterologous element inserted to enhance the level and/or tissue-specificity of a promoter. "Constitutive promoters" refers to those that direct gene expression in all tissues and at all times. "Organ-specific" or "development-specific" promoters as referred to herein are those that direct gene expression almost exclusively in specific organs, such as leaves or seeds, or at specific development stages in an organ, such as in early or late embryogenesis, respectively.

The term "operably linked" refers to nucleic acid sequences on a single nucleic acid molecule which are associated so that the function of one is affected by the other. For example, a promoter is operably linked with a structure gene (i.e., a gene encoding aspartokinase that is lysine-insensitive as given herein) when it is capable of affecting the expression of that structural gene (i.e., that the structural gene is under the transcriptional control of the promoter).

The term "expression", as used herein, is intended to mean the production of the protein product encoded by a gene. More particularly, "expression" refers to the transcription and stable accumulation of the sense (mRNA) or tha antisense RNA derived from the nucleic acid fragment(s) of the invention that, in conjuction with the protein apparatus of the cell, results in altered levels of protein product. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of preventing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

The "3' non-coding sequences" refers to the DNA sequence portion of a gene that contains a polyadenylation signal and any other regulatory signal capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor.

The "translation leader sequence" refers to that DNA sequence portion of a gene between the promoter and coding sequence that is transcribed into RNA and is present in the fully processed mRNA upstream (5') of the translation start codon. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency.

"Mature" protein refers to a post-translationally processed polypeptide without its targeting signal. "Precursor" protein refers to the primary product of translation of mRNA. A "chloroplast targeting signal" is an amino acid sequence which is translated in conjunction with a protein and directs it to the chloroplast. "Chloroplast transit sequence" refers to a nucleotide sequence that encodes a chloroplast targeting signal.

"End-product inhibition" or "feedback inhibition" refers to a biological regulatory mechanism wherein the catalytic activity of an enzyme in a biosynthetic pathway is reversibly inhibited by binding to one or more of the end-products of the pathway when the concentration of the end-product(s) reaches a sufficiently high level, thus slowing the biosynthetic process and preventing over-accumulation of the end-product.

"Transformation" herein refers to the transfer of a foreign gene into the genome of a host organism and its genetically stable inheritance. Examples of methods of plant transformation include Agrobacterium-mediated transformation and particle-accelerated or "gene gun" transformation technology.

"Host cell" means the cell that is transformed with the introduced genetic material.

Isolation of AK-HDH Genes

This invention concerns chimeric genes encoding bi-functional AK-HDH enzymes, wherein both catalytic activities are insensitive to end-product inhibition. Over-expression of feedback-insensitive AK increases flux through the entire pathway of aspartate-derived amino acids even in the presence of high concentrations of the pathway end-products lysine, threonine and methionine. Over-expression of a bifunctional feedback-insensitive AK-HDH enzyme directs the increased flux through the threonine branch of the aspartate-derived amino acid pathway, increasing the rate of threonine biosynthesis.

A number of AK and AK-HDH genes have been isolated and sequenced. These include the thrA gene of *E. coli* (Katinka et al. (1980) *Proc. Natl. Acad. Sci. USA 77*:5730-5733], the metL gene of *E. coli* (Zakin et al. (1983) *J. Biol. Chem. 258*:3028-3031], the lysC gene of *E. coli*[Cassan et al. (1986) *J. Biol. Chem. 261*:1052-1057], and the HOM3 gene of *S. cerevisiae* [Rafalski et al. (1988) *J. Biol. Chem. 263*:2146-2151]. The thrA gene of *E. coli* encodes a bifunctional protein, AKI-HDHI. The AK and HDH activities of this protein is inhibited by threonine. The metL gene of *E. coli* also encodes a bifunctional protein, AKII-HDHII, and both the AK and HDH activities of this protein are insensitive to all pathway end-products. The *E. coli* lysC gene encodes AKIII, which is sensitive to lysine inhibition. The HOM3 gene of yeast encodes an AK which is sensitive to threonine.

Among these microbial genes the *E. coli* metL gene encoding AKII-HDHII is preferred. As indicated above, this gene has been isolated and sequenced. Thus, it can be easily obtained from *E. coli* genomic DNA by a variety of techniques well known to those skilled in the art, for example via PCR using oligonucleotide primers based on the published DNA sequence.

Plant mutants that express lysine-insensitive AK-HDH are known. In barley, lysine plus threonine-resistant mutants bearing mutations in two unlinked genes that result in two different lysine-insensitive AK isoenzymes have been described [Bright et al. (1982) *Nature 299*:278-279, Rognes et al. (1983) *Planta 157*:32-38, Arruda et al. (1984) *Plant Phsiol. 76*:442-446]. In corn, a lysine plus threonine-resistant cell line had AK activity that was less sensitive to lysine inhibition than its parent line [Hibberd et al. (1980) *Planta 148*:183-187]. A subsequently isolated lysine plus threonine-resistant corn mutant is altered at a different genetic locus and also produces lysine-insensitive AK [Diedrick et al. (1990) *Theor. Appl. Genet. 79*:209-215, Dotson et al. (1990) *Planta 182*:546-552]. In tobacco there are two AK enzymes in leaves, one lysine-sensitive and one threonine-sensitive. A lysine plus threonine-resistant tobacco mutant that expressed completely lysine-insensitive AK has been described [Frankard et al. (1991) *Theor. Appl. Genet. 82*:273-282].

These plant mutants could serve as sources of genes encoding feedback-insensitive AK-HDH and used, based on the teachings herein, to increase the accumulation of threonine in the seeds of transformed plants. A partial amino acid sequence of AK from carrot has been reported [Wilson et al. (1991) Plant Physiol. 97:1323:1328]. Using this information a set of degenerate DNA oligonucleotides could be designed, synthesized and used as a hybridization probe to permit the isolation of the carrot AK gene. Recently the carrot AK gene has been isolated

and its nucleotide sequence has been determined [Matthews et al. (1991) U.S.S.N. 07/746,705]. This gene can be isolated based upon its sequence using PCR and used as a hybridization probe to isolate the genes encoding lysine-insensitive AK-HDH described above.

Construction of Chimeric Genes for Expression of AK-HDH in the Seeds of Plants

In order to increase biosynthesis of threonine in seeds, suitable regulatory sequences are provided to create chimeric genes for high level seed-specific expression of the AK-HDH coding region. The replacement of the native regulatory sequences accomplishes two things: 1) any pleiotropic effects that the accumulation of excess free threonine might have on the vegetative growth of plants is prevented because the chimeric gene(s) is not expressed in vegetative tissue of the transformed plants 2) high level expression of the enzyme(s) is obtained in the seeds.

The expression of foreign genes in plants is well-established [De Blaere et al. (1987) *Meth. Enzymol. 143*:277-291]. Proper level of expression of AK-HDH mRNA may require the use of different chimeric genes utilizing different promoters. Such chimeric genes can be transferred into host plants either together in a single expression vector or sequentially using more than one vector. A preferred class of heterologous hosts for the expression of AK-HDH genes are eukaryotic hosts, particularly the cells of higher plants. Particularly preferred among the higher plants and the seeds derived from them are soybean, rapeseed (*Brassica napus, B. campestris*), sunflower (*Helianthus annus*), cotton (*Gossypium hirsutum*), corn, tobacco (*Nicotiana Tubacum*), alfalfa (*Medicago sativa*), wheat (*Triticum sp*), barley (*Hordeum vulgare*), oats (*Avena sativa*, L), sorghum (*Sorghum bicolor*), rice (*Oryza sativa*), and forage grasses. Expression in plants will use regulatory sequences functional in such plants.

The origin of the promoter chosen to drive the expression of the coding sequence is not critical as long as it has sufficient transcriptional activity to accomplish the invention by expressing translatable mRNA for AK-HDH genes in the desired host tissue.

Preferred promoters are those that allow expression of the protein specifically in seeds. This may be especially useful, since seeds are the primary source of vegetable amino acids and also since seed-specific expression will avoid any potential deleterious effect in non-seed organs. Examples of seed-specific promoters include, but are not limited to, the promoters of seed storage proteins. The seed storage proteins are strictly regulated, being expressed almost exclusively in seeds in a highly organ-specific and stage-specific manner [Higgins et al.(1984)

Ann. Rev. Plant Physiol. 35:191-221; Goldberg et al.(1989) *Cell 56*:149-160; Thompson et al. (1989) *BioEssays 10*:108-113]. Moreover, different seed storage proteins may be expressed at different stages of seed development.

There are currently numerous examples for seed-specific expression of seed storage protein genes in transgenic dicotyledonous plants. These include genes from dicotyledonous plants for bean β-phaseolin [Sengupta-Goplalan et al. (1985) *Proc. Natl. Acad. Sci. USA 82*:3320-3324; Hoffman et al. (1988) *Plant Mol. Biol. 11*:717-729], bean lectin [Voelker et al. (1987) *EMBO J. 6*: 3571-3577], soybean lectin [Okamuro et al. (1986) *Proc. Natl. Acad. Sci. USA 83*:8240-8244], soybean kunitz trypsin inhibitor [Perez-Grau et al. (1989) *Plant Cell 1*:095-1109], soybean β-conglycinin [Beachy et al. (1985) *EMBO J. 4*:3047-3053; Barker et al. (1988) *Proc. Natl. Acad. Sci. USA 85*:458-462; Chen et al. (1988) *EMBO J. 7*:297-302; Chen et al. (1989) *Dev. Genet. 10*:112-122; Naito et al. (1988) *Plant Mol. Biol. 11*:109-123], pea vicilin [Higgins et al. (1988) *Plant Mol. Biol. 11*:683-695], pea convicilin [Newbigin et al. (1990) *Planta 180*:461], pea legumin [Shirsat et al. (1989) *Mol. Gen. Genetics 215*:326]; rapeseed napin [Radke et al. (1988) *Theor. Appl. Genet. 75*:685-694] as well as genes from monocotyledonous plants such as for maize 15 kD zein [Hoffman et al. (1987) *EMBO J. 6*:3213-3221; Schernthaner et al. (1988) *EMBO J. 7*:1249-1253; Williamson et al. (1988) *Plant Physiol. 88*:1002-1007], barley β-hordein [Marris et al. (1988) *Plant Mol. Biol. 10*:359-366] and wheat glutenin [Colot et al. (1987) *EMBO J. 6*:3559-3564]. Moreover, promoters of seed-specific genes, operably linked to heterologous coding sequences in chimeric gene constructs, also maintain their temporal and spatial expression pattern in transgenic plants. Such examples include linking either the Phaseolin or Arabidopsis 2S albumin promoters to the Brazil nut 2S albumin coding sequence and expressing such combinations in tobacco, Arabidopsis, or *Brassica napus* [Altenbach et al., (1989) *Plant Mol. Biol. 13*:513-522; Altenbach et al., (1992) Plant *Mol. Biol. 18*:235-245; De Clercq et al., (1990) *Plant Physiol. 94*:970-979], bean lectin and bean β-phaseolin promoters to express luciferase [Riggs et al. (1989) *Plant Sci. 63*:47-57], and wheat glutenin promoters to express chloramphenicol acetyl transferase [Colot et al. (1987) *EMBO J. 6*:3559-3564].

Of particular use in the expression of the nucleic acid fragment of the invention will be the heterologous promoters from several extensively-characterized soybean seed storage protein genes such as those for the Kunitz trypsin inhibitor [Jofuku et al. (1989) *Plant Cell 1*:1079-1093; Perez-Grau et al. (1989) *Plant Cell 1*:1095-1109], glycinin [Nielson et al. (1989) *Plant Cell 1*:313-328], β-conglycinin [Harada et al. (1989) *Plant Cell 1*:415-425]. Promoters

of genes for α'- and β-subunits of soybean β-conglycinin storage protein will be particularly useful in expressing AK-HDH mRNA in the cotyledons at mid- to late-stages of soybean seed development [Beachy et al. (1985) *EMBO J.* 4:3047-3053; Barker et al. (1988) *Proc. Natl. Acad. Sci. USA 85*:458-462; Chen et al. (1988) *EMBO J. 7*:297-302; Chen et al. (1989) *Dev. Genet. 10*:112-122; Naito et al. (1988) *Plant Mol. Biol. 11*:109-123] in transgenic plants, since: a) there is very little position effect on their expression in transgenic seeds, and b) the two promoters show different temporal regulation: the promoter for the α'-subunit gene is expressed a few days before that for the β-subunit gene.

Also of particular use in the expression of the nucleic acid fragments of the invention will be the promoters from several extensively characterized corn seed storage protein genes such as endosperm-specific promoters from the 10 kD zein [Kirihara et al. (1988) *Gene 71*:359-370], the 27 kD zein [Prat et al. (1987) *Gene 52*:51-49; Gallardo et al. (1988) *Plant Sci. 54*:211-281], and the 19 kD zein [Marks et al. (1985) *J. Biol. Chem. 260*:16451-16459]. The relative transcriptional activities of these promoters in corn have been reported [Kodrzyck et al. (1989) *Plant Cell 1*:105-114] providing a basis for choosing a promoter for use in chimeric gene constructs for corn. For expression in corn embryos, a strong embryo-specific promoter, e.g. the promoter from the GLB1 gene [Kriz (1989) *Biochemical Genetics 27*:239-251, Wallace et al. (1991) *Plant Physiol. 95*:973-975] can be used.

It is envisioned that the introduction of enhancers or enhancer-like elements into other promoter constructs will also provide increased levels of primary transcription for AK-HDH genes to accomplish the invention. These would include viral enhancers such as that found in the 35S promoter [Odell et al. (1988) *Plant Mol. Biol. 10*:263-272], enhancers from the opine genes [Fromm et al. (1989) *Plant Cell 1*:977-984], or enhancers from any other source that result in increased transcription when placed into a promoter operably linked to the nucleic acid fragment of the invention.

Of particular importance is the DNA sequence element isolated from the gene for the α'-subunit of β-conglycinin that can confer 40-fold seed-specific enhancement to a constitutive promoter [Chen et al. (1988) *EMBO J.* 7:297-302; Chen et al. (1989) *Dev. Genet. 10*:112-122]. One skilled in the art can readily isolate this element and insert it within the promoter region of any gene in order to obtain seed-specific enhanced expression with the promoter in transgenic plants. Insertion of such an element in any seed-specific gene that is expressed at different times than the β-conglycinin gene will result in expression in transgenic plants for a longer period during seed development.

Any 3' non-coding region capable of providing a polyadenylation signal and other regulatory sequences that may be required for the proper expression of the AK-HDH coding regions can be used to accomplish the invention. This would include the 3' end from any storage protein such as the 3' end of the bean phaseolin gene, the 3' end of the soybean β-conglycinin gene, the 3' end from viral genes such as the 3' end of the 35S or the 19S cauliflower mosaic virus transcripts, the 3' end from the opine synthesis genes, the 3' ends of ribulose 1,5-bisphosphate carboxylase or chlorophyll a/b binding protein, or 3' end sequences from any source such that the sequence employed provides the necessary regulatory information within its nucleic acid sequence to result in the proper expression of the promoter/coding region combination to which it is operably linked. There are numerous examples in the art that teach the usefulness of different 3' non-coding regions [for example, see Ingelbrecht et al. (1989) *Plant Cell 1*:671-680].

DNA sequences coding for intracellular localization sequences may be added to the AK-HDH coding sequence if required for the proper expression of the proteins to accomplish the invention. Plant amino acid biosynthetic enzymes are known to be localized in the chloroplasts and therefore are synthesized with a chloroplast targeting signal. Bacterial proteins such as *E. coli* AKII-HDHII have no such signal. A chloroplast transit sequence could, therefore, be fused to the coding sequence. Preferred chloroplast transit sequences are those of the small subunit of ribulose 1,5-bisphosphate carboxylase, e.g. from soybean [Berry-Lowe et al. (1982) *J. Mol. Appl. Genet. 1*:483-498] for use in dicotyledonous plants and from corn [Lebrun et al. (1987) *Nucleic Acids Res.* 15:4360] for use in monocotyledonous plants.

Introduction of Chimeric Genes into Plants

Various methods of introducing a DNA sequence into eukaryotic cells (i.e., of transformation) of higher plants are available to those skilled in the art (see EPO publications 0 295 959 A2 and 0 138 341 A1). Such methods include those based on transformation vectors utilizing the Ti and Ri plasmids of *Agrobacterium spp.* It is particularly preferred to use the binary type of these vectors. Ti-derived vectors transform a wide variety of higher plants, including monocotyledonous and dicotyledonous plants, such as soybean, cotton and rape [Pacciotti et al. (1985) *Bio/Technology 3*:241; Byrne et al. (1987) *Plant Cell, Tissue and Organ Culture 8*:3; Sukhapinda et al. (1987) *Plant Mol. Biol. 8*:209-216; Lorz et al. (1985) *Mol. Gen. Genet. 199*:178; Potrykus (1985) *Mol. Gen. Genet. 199*:183].

Other transformation methods are available to those skilled in the art, such as direct uptake of foreign DNA constructs [see EPO publication 0 295 959 A2], techniques of electroporation [see Fromm et al. (1986) *Nature* (London) *319*:791]

or high-velocity ballistic bombardment with metal particles coated with the nucleic acid constructs [see Kline et al. (1987) *Nature* (London) *327*:70, and see U.S. Pat. No. 4,945,050]. Once transformed, the cells can be regenerated by those skilled in the art.

Of particular relevance are the recently described methods to transform foreign genes into commercially important crops, such as rapeseed [see De Block et al. (1989) *Plant Physiol. 91*:694-701], sunflower [Everett et al. (1987) *Bio/Technology 5*:1201], soybean [McCabe et al. (1988) *Bio/Technology 6*:923; Hinchee et al. (1988) *Bio/Technology 6*:915; Chee et al. (1989) *Plant Physiol. 91*:1212-1218; Christou et al. (1989) *Proc. Natl. Acad. Sci USA 86*:7500-7504; EPO Publication 0 301 749 A2], and corn [Gordon-Kamm et al. (1990) *Plant Cell 2*:603-618; Fromm et al. (1990) *Biotechnology 8*:833-839].

Expression of Chimeric Genes in Transformed Plants

To analyze for expression of the chimeric AK-HDH gene in seeds and for the consequences of expression on the amino acid content in the seeds, a seed meal can be prepared by any of a number of suitable methods known to those skilled in the art. The seed meal can be partially or completely defatted, via hexane extraction for example, if desired. Protein extracts can be prepared from the meal and analyzed for AK or HDH enzyme activities. Alternatively the presence of any of the proteins can be tested for immunologically by methods well-known to those skilled in the art. To measure free amino acid composition of the seeds, free amino acids can be extracted from the meal and analyzed by methods known to those skilled in the art [Bieleski et al. (1966) Anal. *Biochem. 17*:278-293]. Amino acid composition can then be determined using any commercially available amino acid analyzer. To measure total amino acid composition of the seeds, meal containing both protein-bound and free amino acids can be acid hydrolyzed to release the protein-bound amino acids and the composition can then be determined using any commercially available amino acid analyzer. Seeds expressing the AK-HDH protein and with higher threonine content than the wild type seeds can thus be identified and propagated.

EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and

scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLE 1

Isolation of the *E. coli* metL Gene and Over-Expression if AKII-HDHIII in *E. coli*

The metL gene of *E. coli* encodes a bifunctional protein, AKII-HDHIII; the AK and HDH activities of this enzyme are insensitive to all pathway end-products. The metL gene of *E. coli* has been isolated and sequenced previously [Zakin et al. (1983) *J. Biol. Chem. 258*:3028-3031]. For the present invention a DNA fragment containing the metL gene was isolated and modified from *E. coli* genomic DNA obtained from strain LE392 using PCR. The following PCR primers were designed and synthesized:

CF23 = SEQ ID NO:1:
5'-GAAACCATGG CCAGTGTGAT TGCGCAGGCA-3'

CF24 = SEQ ID NO:2:
5'-GAAAGGTACC TTACAACAAC TGTGCCAGC-3'

These primers add an Nco I site which includes a translation initiation codon at the amino terminus of the AKII-HDHIII protein. In order to add the restriction site and additional codon, GCC coding for alanine, was also added to the amino terminus of the protein. The primers also add a Kpn I site immediately following the translation stop codon.

PCR was performed using a Perkin-Elmer Cetus kit according to the instructions of the vendor on a thermocycler manufactured by the same company. The primers were at a concentration of 10 µM and the thermocycling conditions were:

94° 1 min, 50° 2 min, 72° 8 min for 10 cycles followed by
94° 1 min, 72° 8 min for 30 cycles.

Reactions with four different concentrations of template DNA all yielded the expected 2.4 kb DNA fragment, along with several other smaller fragments. The four PCR reaction mixes were pooled, digested with Nco I and Kpn I and the 2.4 kb fragments were purified and isolated from an agarose gel.

To achieve high level expression of the metL gene in *E. coli* an expression vector base upon pET-3a [Rosenberg et al. (1987) *Gene* 56:125-135] which employs the bacteriophage T7 RNA polymerase/T7 promoter system was

constructed. First the EcoR I and Hind III sites in pET-3a were destroyed at their original positions by cutting, filling the ends using the Klenow fragment of DNA polymerase and religating. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG, creating plasmid pBT430. This plasmid was further modified by additon of a Kpn I site downstream of the Nco I site using oligonucleotide adaptors.

The 2.4 kb Nco I and Kpn I metL fragments described above were inserted into the modified pBT430 expression vector cut with Nco I and Kpn I. DNA was isolated from 8 clones carrying the 2.4 kb fragment in the expression vector and transformed into the expression host strain BL21(DE3).

Cultures were grown in TB medium containing ampicillin (100 mg/L) at 37°C overnight. The cells were collected by centrifugation and resuspended in 1/25th the original culture volume in 50 mM NaCl; 50 mM Tris-Cl, pH 7.5; 1 mM EDTA, and frozen at -20°C, thawed at 37°C and sonicated, in an ice-water bath, to lyse the cells. The lysate was centrifuged at 4°C for 5 min at 12,000 rpm. The supernatant was removed and the pellet was resuspended in the above buffer.

The supernatant fractions were assayed for HDH enzyme activities to identify clones expressing functional proteins. HDH activity was assayed as shown below:

HDH ASSAY

Stock solutions	1.0 mL	0.20 mL	Final conc
0.2 M KPO_4, pH 7.0	500 µL	100 µL	100 mM
3.7 M KCl	270 µL	54 µL	1.0 M
0.5 M EDTA	20 µL	4 µL	10 mM
1.0 M $MgCl_2$	10 µL	2 µL	10 mM
2 mM NADPH	100 µL	20 µL	0.20 mM

Make Mixture of above reagents with amounts multiplied by number of assays. Use 0.9 mL of mix for 1ml assay; 180 µl of mix for 0.2 mL assay in microtiter dish.

Add			
1.0M ASA in 1.0N HCl	1μL	0.2μL	1.0mM
to 1/2 the assay mix; remaining 1/2 lacks ASA to serve as blank			
enzyme extract	10-100 μL	2-20 μL	
H_2O	to 1.0 mL	to 0.20 mL	

Add enzyme extract last to start reaction. Incubate at ~30°C; monitor NADPH oxidation at 340 nM. 1 unit oxidizes 1 μmol NADPH/min at 30°C in the 1 ml reaction.

Four of eight extracts showed HDH activity well above the untransformed *E. coli* host. These four were then assayed for AK activity. AK activity was assayed as shown below:

AK ASSAY

Assay mix (for 12 X 1.0mL or 48 X 0.25mL assays):

- 2.5 mL H_2O
- 2.0 mL 4M KOH
- 2.0 mL 4M NH_2OH-HCl
- 1.0 mL 1M Tris-HCl pH 8.0
- 0.5 mL 0.2M ATP (121 mg/mL in 0.2M NaOH)
- 50 μL 1M $MgSO_4$

pH of assay mix should be 7-8

Each 1.5 mL eppendorf assay tube contains:

	MACRO assay	micro assay
assay mix	0.64 mL	0.16 mL
0.2M L-Aspartate	0.04 mL	0.01 mL
extract	5-120 μL	1-30 μL
H_2O to total vol.	0.8 mL	0.2 mL
Assay tubes are incubated at 30°C for 30-60 min		
Add to develop color;		
$FeCl_3$ reagent	0.4 mL	0.1 mL
$FeCl_3$ reagent is:	10% w/v $FeCl_3$	50 g
	3.3% TCA	15.5 g
	0.7% HCl	35 mls HCl
		H_2O to 500 mls

Spin for 2 min in eppendorf centrifuge tube.

Read OD at 540 nm.

Two extracts also had high levels of AK enzyme activity. These two extracts were then tested for inhibition of AK or HDH activity by the pathway end-products, lys, thr and met. Neither the AK nor the HDH activity of the extract from clone 5 was inhibited by 30 mM concentrations of any of the end-products.

The supernatant and pellet fractions of several of the extracts were also analyzed by SDS polyacrylamide gel electrophoresis. In the extract from clone 5, the major protein visible by Coomassie blue staining in both the pellet and supernatant fractions had a molecular weight of about 85 kd, the expected size for AKII-HDHII. The metL gene in plasmid pBT718 (Figure 1) from clone 5 was used for all subsequent work. AKII-HDHII protein derived from clone 5 was sent to Hazelton Research Facility (310 Swampridge Road, Denver, PA 17517) to have rabbit antibodies raised against the protein.

EXAMPLE 2

Construction of Chimeric Genes for Expression *E. coli* metL in the Embryo and Endosperm of Transformed Corn

The following chimeric genes were made for transformation into corn:

globulin 1 promoter/mcts/metL/globulin 1 3' region
glutelin 2 promoter/mcts/metL/NOS 3' region

The globulin 1 promoter and 3' sequences were isolated from a Clontech corn genomic DNA library using oligonucleotide probes based on the published sequence of the globulin 1 gene [Kriz et al. (1989) *Plant Physiol. 91*:636]. The cloned segment includes the promoter fragment extending 1078 nucleotides upstream from the ATG translation start codon, the entire globulin coding sequence including introns and the 3' sequence extending 803 bases from the translational stop. To allow replacement of the globulin 1 coding sequence with other coding sequences an Nco I site was introduced at the ATG start codon, and Kpn I and Xba I sites were introduced following the translational stop codon via PCR to create vector pCC50. There is a second Nco I site within the globulin 1 promoter fragment. The globulin 1 gene cassette is flanked by Hind III sites.

The glutelin 2 promoter was cloned from corn genomic DNA using PCR with primers based on the published sequence [Reina et al. (1990) *Nucleic Acids Res. 18*:6426-6426]. The promoter fragment includes 1020 nucleotides upstream from the ATG translation start codon. An Nco I site was introduced via PCR at the ATG start site to allow for direct translational fusions. A BamH I site was introduced on the 5' end of the promoter. The 1.02 kb BamH I to Nco I promoter fragment was cloned into the BamH I to Nco I sites of a previously constructed

plant expression vector replacing the 35S promoter to create vector pML90. This vector contains the glutelin 2 promoter linked to the GUS coding region and the NOS 3'.

Plant amino acid biosynthetic enzymes are known to be localized in the chloroplasts and therefore are synthesized with a chloroplast targeting signal. Bacterial proteins have no such signal. A chloroplast transit sequence (cts) was therefore fused to the *E. coli* metL coding sequence in the chimeric genes described below. For corn the cts used was based on the the cts of the small subunit of ribulose 1,5-bisphosphate carboxylase from corn [Lebrun et al. (1987) *Nucleic Acids Res. 15*:4360] and is designated mcts.

Oligonucleotides SEQ ID NO:3 and SEQ ID NO:4, which encode the carboxy terminal part of the corn chloroplast targeting signal, were annealed, resulting in Xba I and Nco I compatible ends, purified via polyacrylamide gel electrophoresis, and inserted into Xba I plus Nco I digested pBT718 (Figure 1). The insertion of the correct sequence was verified by DNA sequencing yielding pBT725. To complete the corn chloroplast targeting signal, the amino terminal part of the corn chloroplast targeting signal from a previous construct, pBT580, was inserted.

The plasmid pBT580 was constructed as follows: oligonucleotides SEQ ID NO:3 and SEQ ID NO:4 (above) were annealed, resulting in Xba I and Nco I compatible ends, purified via polyacrylamide gel electrophoresis, and inserted into an Xba I plus Nco I digested plasmid containing the *E. coli* lysC gene, thus fusing the carboxy terminal part of the corn chloroplast targeting signal to the AKIII protein, and destroying the Nco I site. The insertion of the correct sequence was verified by DNA sequencing yielding pBT556. Oligonucleotides SEQ ID NO:5 and SEQ ID NO:6, which encode the middle part of the chloroplast targeting signal, were annealed, resulting in Bgl II and Xba I compatible ends, purified via polyacrylamide gel electrophoresis, and inserted into Bgl II and Xba I digested pBT556. The insertion of the correct sequence was verified by DNA sequencing yielding pBT557. Oligonucleotides SEQ ID NO:7 and SEQ ID NO:8, which encode the amino terminal part of the chloroplast targeting signal, were annealed, resulting in Nco I and Afl II compatible ends, purified via polyacrylamide gel electrophoresis, and inserted into Nco I and Afl II digested pBT557. The insertion of the correct sequence was verified by DNA sequencing yielding pBT558. Thus the complete mcts was fused to the lysC gene.

The mcts/lysC-M4 coding sequence was isolated from plasmid pBT558 and inserted into Nco I plus Sma I digested pML90 (above) creating plasmid pBT580. pBT580 was digested with BamH I and Xba I yielding a 1.14 kb fragment

containing the glutelin 2 promoter plus the amino terminal part of the corn chloroplast targeting signal. This fragment was inserted into pBT725 digested with Bgl II and Xba I, creating pBT726 (Figure 2) wherein the complete mcts was fused to the metL coding sequence.

To construct the chimeric gene:
globulin 1 promoter/mcts/metL/globulin 1 3' region
the 2.6 kb Nco I to Kpn I fragment containing the mcts/metL coding sequence was isolated from plasmid pBT726 and inserted into Nco I (partial digest) plus Kpn I digested pCC50 creating plasmid pBT727 (Figure 3).

To construct the chimeric gene:
glutelin 2 promoter/mcts/metL/NOS 3' region
the 2.6 kb Nco I to Kpn I fragment containing the mcts/metL coding sequence was isolated from plasmid pBT726 and inserted into pML90 (described above) digested with the same enzymes creating plasmid pBT728 (Figure 4).

EXAMPLE 3

Transformation of Corn with Chimeric Genes for Expression of *E. coli* metL in the Embryo and Endosperm

Corn was transformed with the chimeric genes:

globulin 1 promoter/mcts/metL/globulin 1 3' region

or

glutelin 2 promoter/mcts/metL/NOS 3' region

The bacterial bar gene from *Streptomyces hygroscopicus* that confers resistance to the herbicide glufosinate [Thompson et al. (1987) The *EMBO Journal* 6:2519-2523] was used as the selectable marker for corn transformation. The bar gene had its translation codon changed from GTG to ATG for proper translation initiation in plants [De Block et al. (1987) The *EMBO Journal* 6:2513-2518]. The bar gene was driven by the 35S promoter from Cauliflower Mosaic Virus and uses the termination and polyadenylation signal from the octopine synthase gene from *Agrobacterium tumefaciens*.

Embryogenic callus cultures were initiated from immature embryos (about 1.0 to 1.5 mm) dissected from kernels of a corn line bred for giving a "type II callus" tissue culture response. The embryos were dissected 10 to 12 d after pollination and were placed with the axis-side down and in contact with agarose-solidified N6 medium [Chu et al. (1974) *Sci Sin 18*:659-668] supplemented with 1.0 mg/L 2,4-D (N6-1.0). The embryos were kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryos and somatic embryos borne on suspensor structures proliferated from

the scutellum of the immature embryos. Clonal embryogenic calli isolated from individual embryos were identified and sub-cultured on N6-1.0 medium every 2 to 3 weeks.

The particle bombardment method was used to transfer genes to the callus culture cells. A Biolistic™ PDS-1000/He (BioRAD Laboratories, Hercules, CA) was used for these experiments.

Circular plasmid DNA or DNA which had been linearized by restriction endonuclease digestion was precipitated onto the surface of gold particles. DNA from two or three different plasmids, one containing the selectable marker for corn transformation, and one containing the chimeric gene for increased threonine accumulation in seeds were co-precipitated. To accomplish this 2.5 μg of each DNA (in water at a concentration of about 1 mg/mL) was added to 25 μL of gold particles (average diameter of 1.0 μm) suspended in water (60 mg of gold per mL). Calcium chloride (25 μL of a 2.5 M solution) and spermidine (10 μL of a 0.1 M solution) were then added to the gold-DNA suspension as the tube was vortexing for 3 min. The gold particles were centrifuged in a microfuge for 1 sec and the supernatant removed. The gold particles were then resuspended in 1 mL of absolute ethanol, were centrifuged again and the supernatant removed. Finally, the gold particles were resuspended in 25 μL of absolute ethanol and sonicated twice for one sec. Five μL of the DNA-coated gold particles were then loaded on each macro carrier disk and the ethanol was allowed to evaporate away leaving the DNA-covered gold particles dried onto the disk.

Embryogenic callus was arranged in a circular area of about 4 cm in diameter in the center of a 100 X 20 mm petri dish containing N6-1.0 medium supplemented with 0.25M sorbitol and 0.25M mannitol. The tissue was placed on this medium for 4-6 h prior to bombardment as a pretreatment and remained on the medium during the bombardment procedure. At the end of the 4-6 h pretreatment period, the petri dish containing the tissue was placed in the chamber of the PDS-1000/He. The air in the chamber was then evacuated to a vacuum of 28-29 inch of Hg. The macrocarrier was accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1080-1100 psi. The tissue was placed approximately 8 cm from the stopping screen. Five to seven plates of tissue were bombarded with the DNA-coated gold particles. Following bombardment, the callus tissue was transferred to N6-1.0 medium without supplemental sorbitol or mannitol.

Within 3-5 days after bombardment the tissue was transferred to selective medium, N6-1.0 medium that contained 2 mg/L bialaphos. All tissue was transferred to fresh N6-1.0 medium supplemented with bialaphos every 2 weeks.

After 6-12 weeks clones of actively growing callus were identified. Callus was then transferred to an MS-based medium that promotes plant regeneration.

EXAMPLE 4

Construction of a Chimeric Gene for Expression *E. coli* metL in the Seed of Transformed Soybean

The following chimeric gene was made for transformation into soybean:

phaseolin 5' region/cts/metL/phaseolin 3' region

A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the β subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* [Doyle et al. (1986) *J. Biol. Chem. 261*:9228-9238] was used for expression in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), Sma I, Kpn I and Xba I. The entire cassette is flanked by Hind III sites.

To construct the chimeric gene:

phaseolin 5' region/cts/metL/phaseolin 3' region

the 2.4 kb Nco I to Kpn I fragment containing the metL coding sequence was isolated from plasmid pBT718 (above) and inserted into a pUC18 vector carrying the seed expression cassette also digested with Nco I plus Kpn I creating plasmid pBT733 (Figure 5).

Plant amino acid biosynthetic enzymes are known to be localized in the chloroplasts and therefore are synthesized with a chloroplast targeting signal. The bacterial protein AKII-HDHII has no such signal. A chloroplast transit sequence (cts) was therefore fused to the *E. coli* metL coding sequence in the chimeric gene. The cts used, SEQ ID NO:9:, was equivalent to the the cts of the small subunit of ribulose 1,5-bisphosphate carboxylase from soybean [Berry-Lowe et al. (1982) *J. Mol. Appl. Genet. 1*:483-498]. The cts was flanked with Nco I and inserted into the metL seed expression cassette of pBT733 creating plasmid pBT766 (Figure 6).

EXAMPLE 5

Transformation of Soybean with a Phaseolin/cts/metL Chimeric Gene

To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, were cultured in the light or dark at 26°C on an agar medium (SB1 or SB2) for 6-10 weeks. Somatic embryos, which produced secondary embryos were excised and placed into a liquid medium (SB55). After repeated selection for clusters of somatic

embryos which multiplied as early, globular staged embryos, the suspensions were maintained as described below.

Soybean embryogenic suspension cultures were maintained in 35 mL liquid media (SB55) on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures were subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures were transformed by the method of particle gun bombardment [Kline et al. (1987) *Nature* (London) *327*:70, U.S. Patent No. 4,945,050]. A Du Pont Biolistic™ PDS1000/HE instrument (helium retrofit) was used for these transformations.

The selectable marker gene for soybean transformation was a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus [Odell et al.(1985) *Nature 313*:810-812], the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*) [Gritz et al.(1983) *Gene 25*:179-188] and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette, phaseolin 5' region/cts/metL/phaseolin 3' region, (Example 4) was isolated as an approximately 4.5 kb Hind III fragment from pBT766 (Figure 6). This fragment was inserted into a unique Hind III site of the vector carrying the marker gene creating plasmid pBT767.

To 50 μL of a 60 mg/mL 1 μm gold particle suspension was added (in order); 5 μL DNA (1 μg/μL), 20 μl spermidine (0.1 M), and 50 μL $CaCl_2$ (2.5 M). The particle preparation was agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles were then washed once in 400 μL 70% ethanol and resuspended in 40 μL of anhydrous ethanol. The DNA/particle suspension was sonicated three times for one second each. Five μL of the DNA-coated gold particles were then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture was placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure was set at 1100 psi and the chamber was evacuated to a vacuum of 28 inches mercury. The tissue was placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue was divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment the liquid media was exchanged with fresh SB55, and eleven to twelve days post bombardment with fresh SB55 containing 50 mg/mL hygromycin. The selective media was refreshed weekly.

Seven to eight weeks post bombardment, green, transformed tissue was observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue was removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line was treated as an independent transformation event. These suspensions could then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

Media:

SB55 Stock Solutions (grams per liter):

MS Sulfate 100X Stock

$MgSO_4\ 7H_2O$	37.0
$MnSO_4\ H_2O$	1.69
$ZnSO_4\ 7H_2O$	0.86
$CuSO_4\ 5H_2O$	0.0025

MS Halides 100X Stock

$CaCl_2\ 2H_2O$	44.0
KI	0.083
$CoCl_2\ 6H_20$	0.00125

MS P,B,Mo 100X Stock

KH_2PO_4	17.0
H_3BO_3	0.62
$Na_2MoO_4\ 2H_2O$	0.025

MS FeEDTA 100X Stock

Na_2EDTA	3.724
$FeSO_4\ 7H_2O$	2.784

B5 Vitimin Stock

10 g m-inositol
100 mg nicotinic acid
100 mg pyridoxine HCl
1 g thiamine

SB55 (per liter)

10 mL each MS stocks
1 mL B5 Vitamin stock
0.8 g NH_4NO_3
3.033 g KNO_3
1 mL 2,4-D (10 mg/mL stock)
60 g sucrose
0.667 g asparagine
pH 5.7

<u>SB103</u> (per liter)
MS Salts
6% maltose
750 mg $MgCl_2$
0.2% Gelrite
pH 5.7

<u>SB1</u> (per liter)
MS Salts
B5 Vitamins
0.175 M glucose
20 mg 2,4-D
0.8% agar
pH 5.8

<u>SB2</u>
same as SB1 except 40 mg/L 2,4-D

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: E. I. DU PONT DE NEMOURS AND COMPANY

(ii) TITLE OF INVENTION: CHIMERIC GENES AND METHOD FOR INCREASING THE THREONINE CONTENT OF THE SEEDS OF PLANTS

(iii) NUMBER OF SEQUENCES: 9

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
(B) STREET: 1007 MARKET STREET
(C) CITY: WILMINGTON
(D) STATE: DELAWARE
(E) COUNTRY: U.S.A.
(F) ZIP: 19898

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Microsoft Word 6.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: BARBARA C. SIEGELL
(B) REGISTRATION NUMBER: 30,684
(C) REFERENCE/DOCKET NUMBER: BB-1063-A

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 302-992-4931
(B) TELEFAX: 302-773-0164
(C) TELEX: 835420

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAAACCATGG CCAGTGTGAT TGCGCAGGCA 30

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAAAGGTACC TTACAACAAC TGTGCCAGC 29

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTAGAAGCCT CGGCAACGTC AGCAACGGCG GAAGAATCCG GTG 43

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CATGCACCGG ATTCTTCCGC CGTTGCTGAC GTTGCCGAGG CTT 43

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GATCCCATGG CGCCCCTTAA GTCCACCGCC AGCCTCCCCG TCGCCCGCCG CTCCT 55

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTAGAGGAGC GGCGGGCGAC GGGGAGGCTG GCGGTGGACT TAAGGGGCGC CATGG 55

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CATGGCGCCC ACCGTGATGA TGGCCTCGTC GGCCACCGCC GTCGCTCCGT TCCAGGGGC 59

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TTAAGCCCCT GGAACGGAGC GACGGCGGTG GCCGACGAGG CCATCATCAC GGTGGGCGC 59

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 174 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCATGGCTTC CTCAATGATC TCCTCCCCAG CTGTTACCAC CGTCAACCGT GCCGGTGCCG 60

GCATGGTTGC TCCATTCACC GGCCTCAAAA GCATGGCTGG CTTCCCCACG AGGAAGACCA 120

ACAATGACAT TACCTCCATT GCTAGCAACG GTGGAAGAGT ACAATGTGCC ATGG 174

CLAIMS

What is claimed is:

1. A chimeric gene wherein a nucleic acid fragment encoding a bi-functional protein with aspartokinase and homoserine dehydrogenase activities, both of which are substantially insensitive to end-product inhibition, is operably linked to a plant chloroplast transit sequence and to a seed-specific regulatory sequence.

2. The chimeric gene of Claim 1 wherein the nucleic acid fragment comprises the *E. coli* metL gene.

3. A plant comprising in its genome the chimeric gene of Claim 1.

4. Seeds obtained from the plant of Claim 3.

5. A method for increasing the threonine content of the seeds of plants above the level found in seeds of untransformed plants comprising:

   (a) transforming plant cells with the chimeric gene of Claim 1;

   (b) growing fertile mature plants from the transformed plant cells obtained from step (a) under conditions suitable to obtain seeds; and

   (c) selecting from the progeny seed of step (b) for those seeds containing increased levels of threonine compared to untransformed seeds.

6. A plant produced by the method of Claim 5.

7. A plant comprising in its genome the chimeric gene of Claim 2.

8. Seeds obtained from the plant of Claim 7.

9. A method for increasing the threonine content of the seeds of plants comprising:

   (a) transforming plant cells with the chimeric gene of Claim 2;

   (b) growing fertile mature plants from the transformed plant cells obtained from step (a) under conditions suitable to obtain seeds; and

   (c) selecting from the progeny seed of step (b) for those seeds containing increased levels of threonine compared to untransformed seeds.

10. A plant produced by the method of Claim 9.

FIG. 1

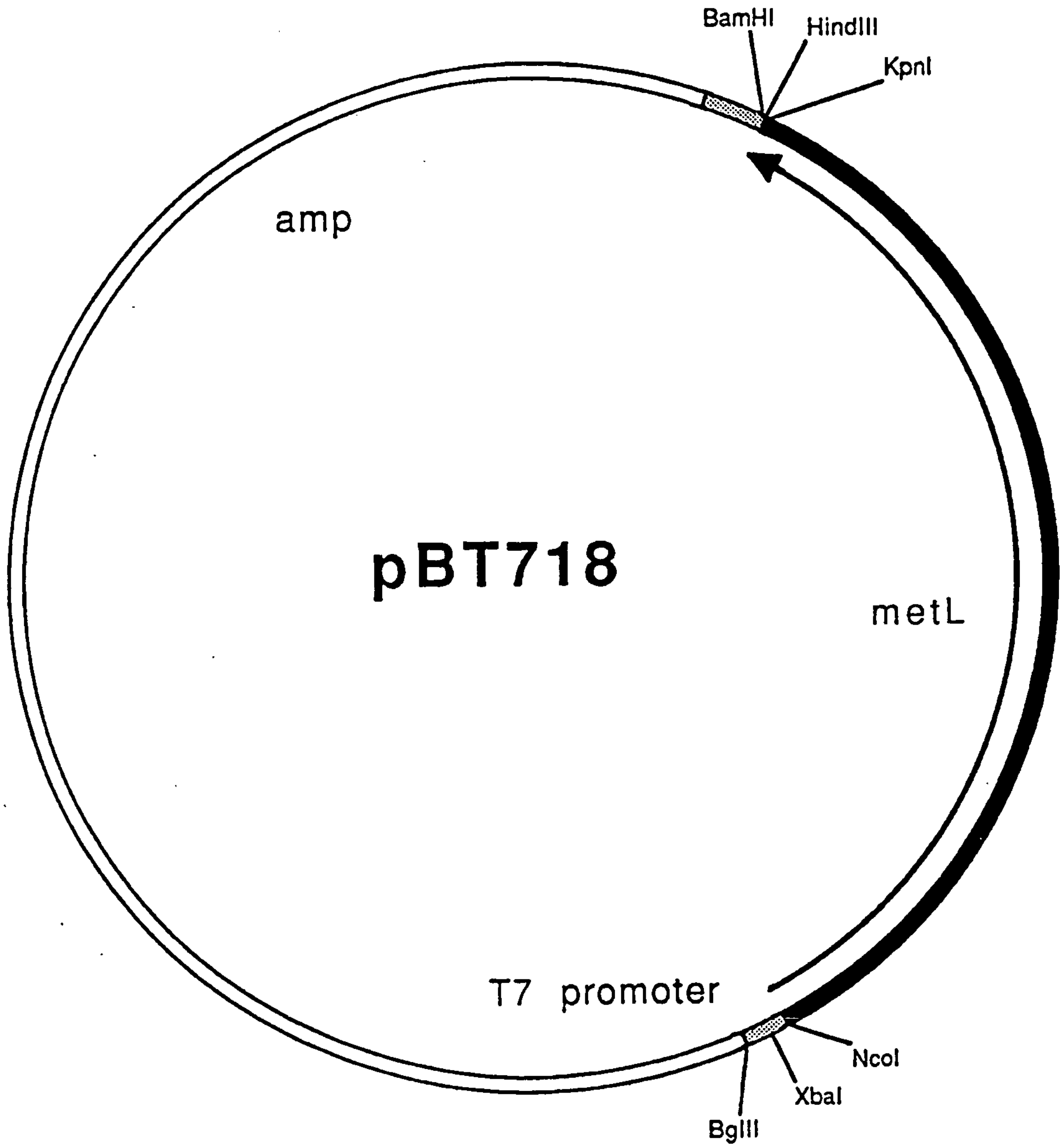

BamHI
HindIII
KpnI
amp
pBT718
metL
T7 promoter
NcoI
XbaI
BglII

FIG. 2

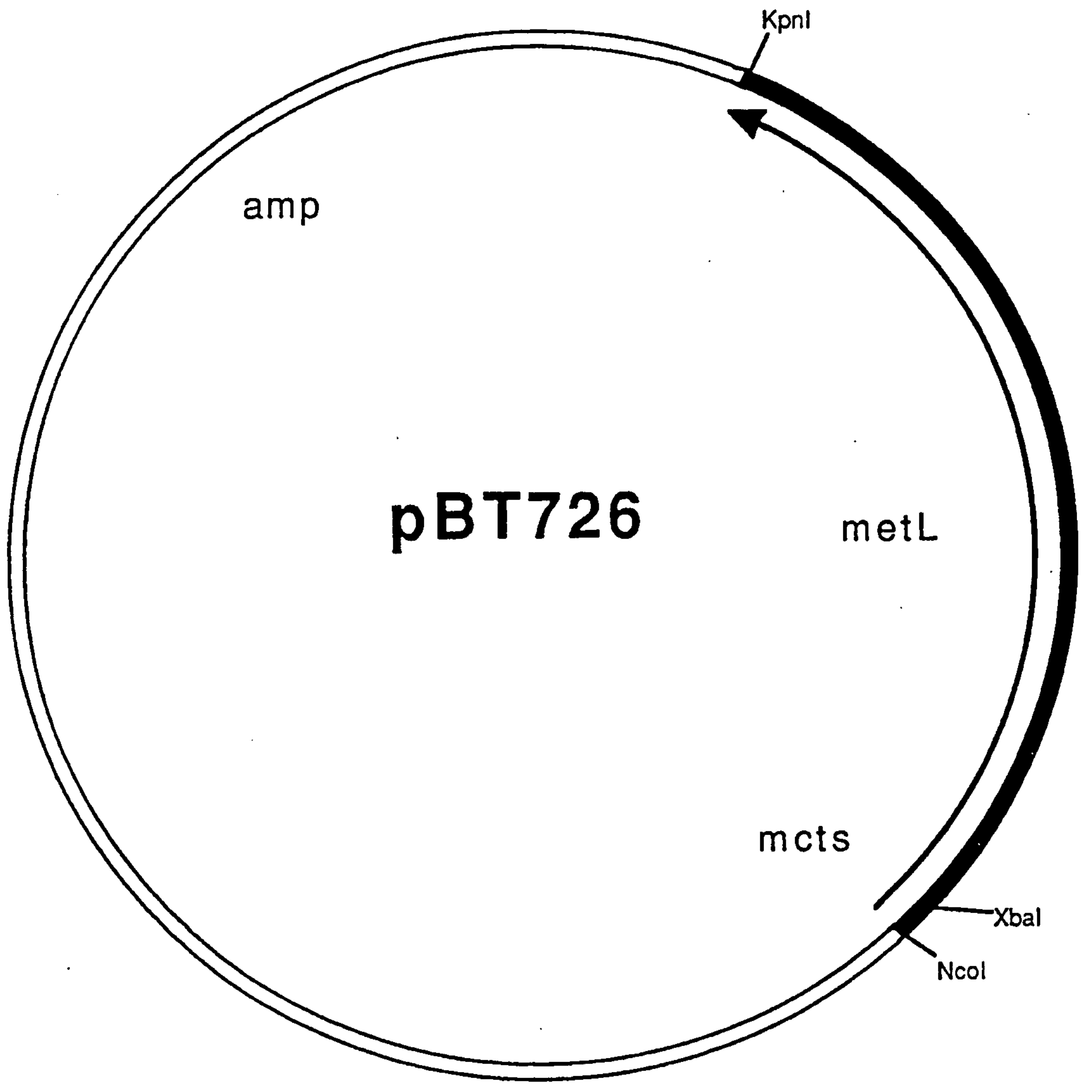


FIG. 3

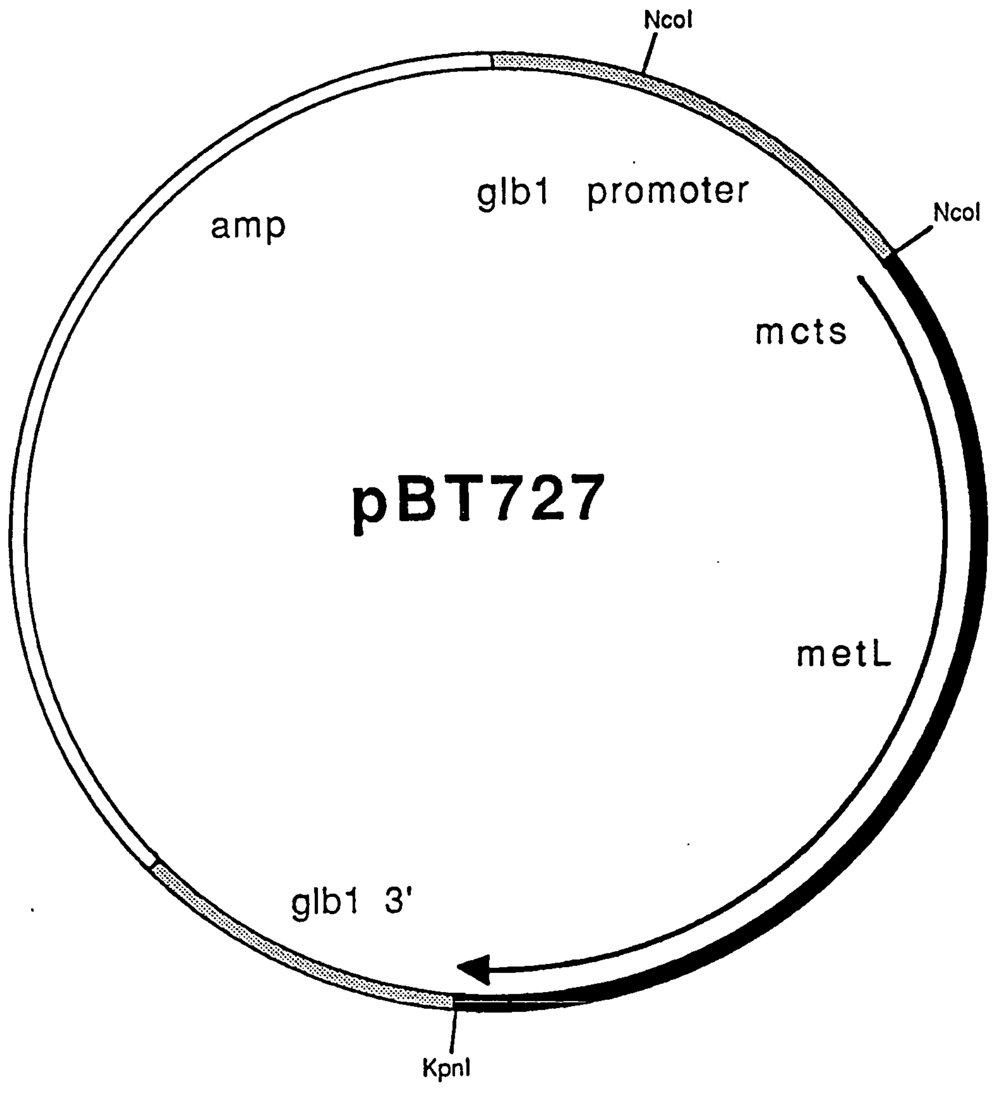


FIG. 4

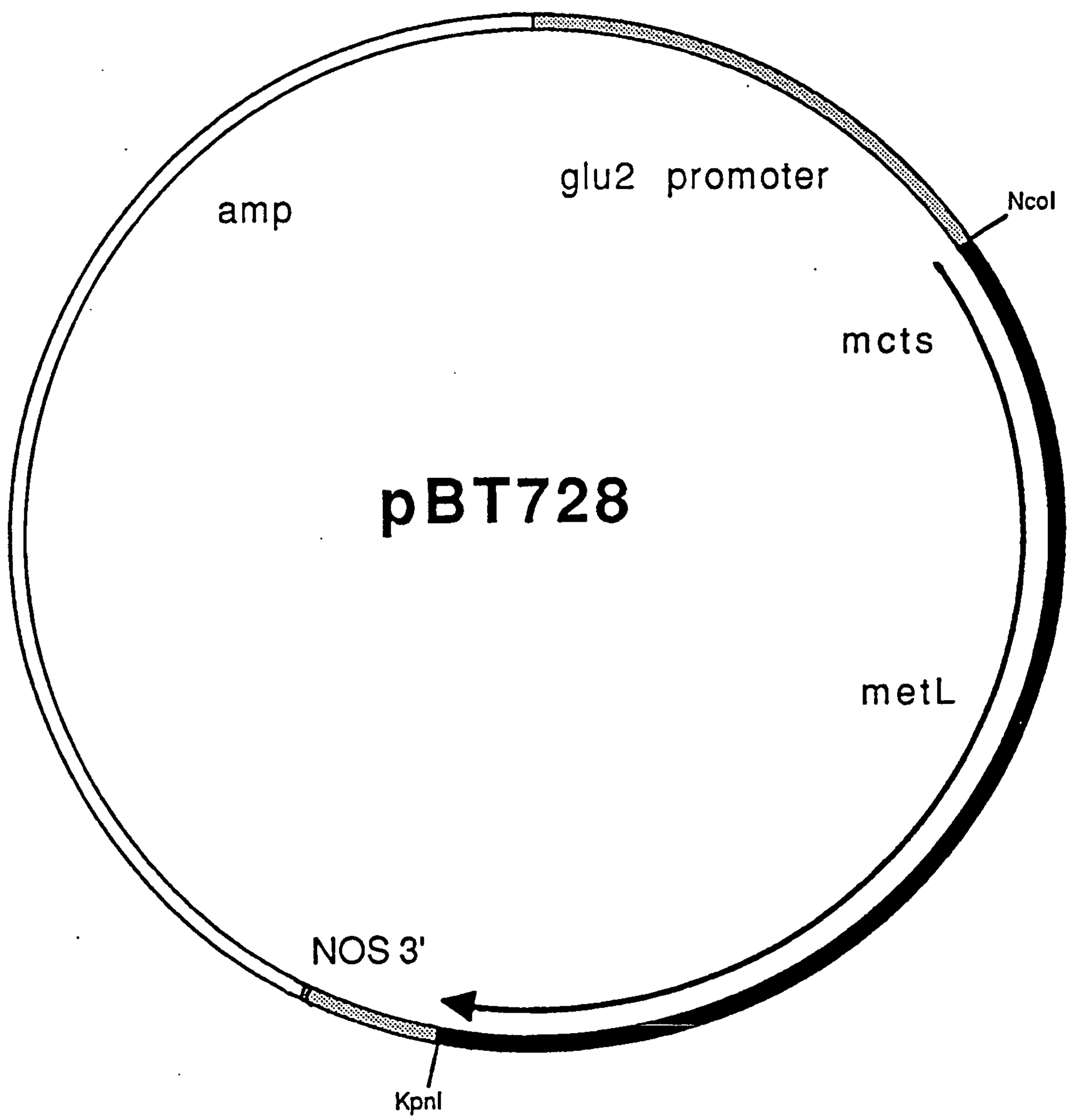


FIG. 5

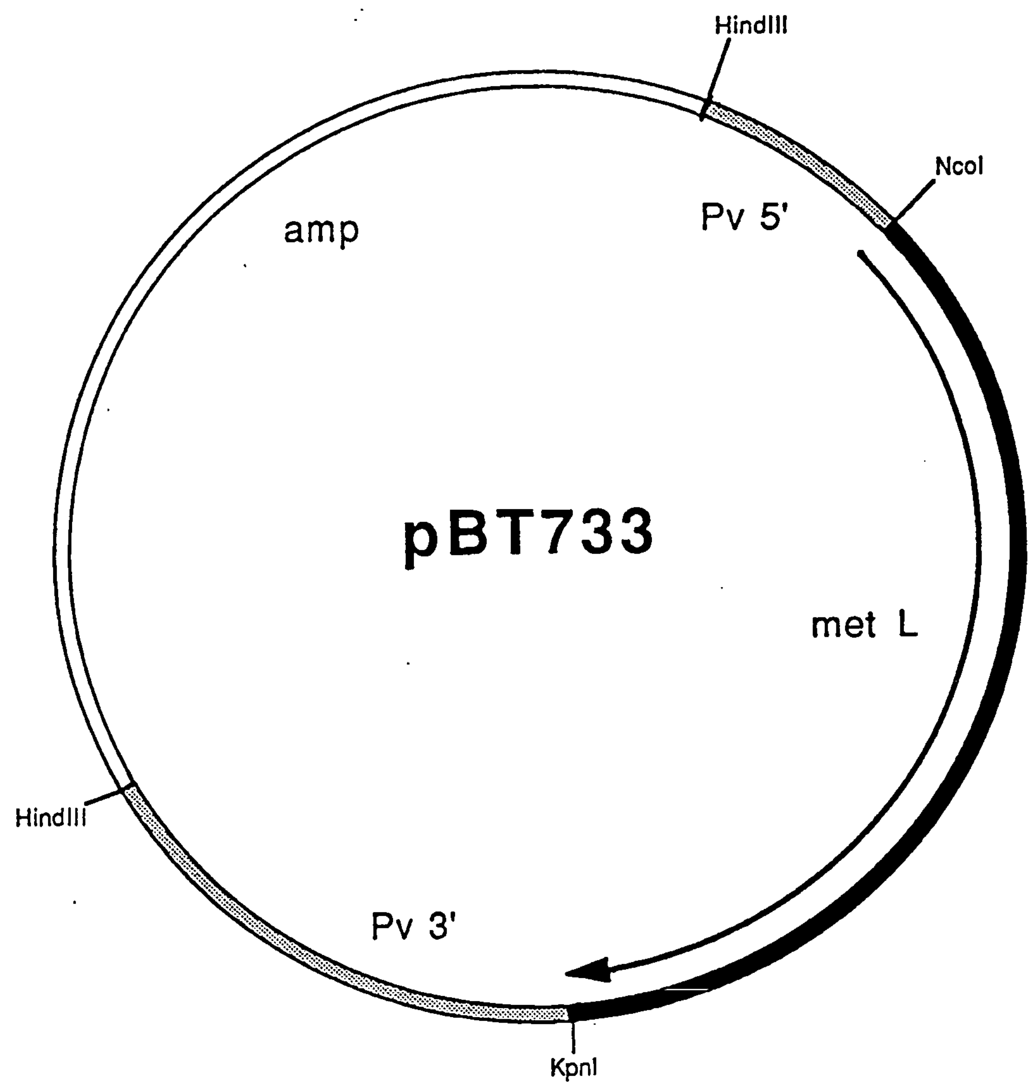
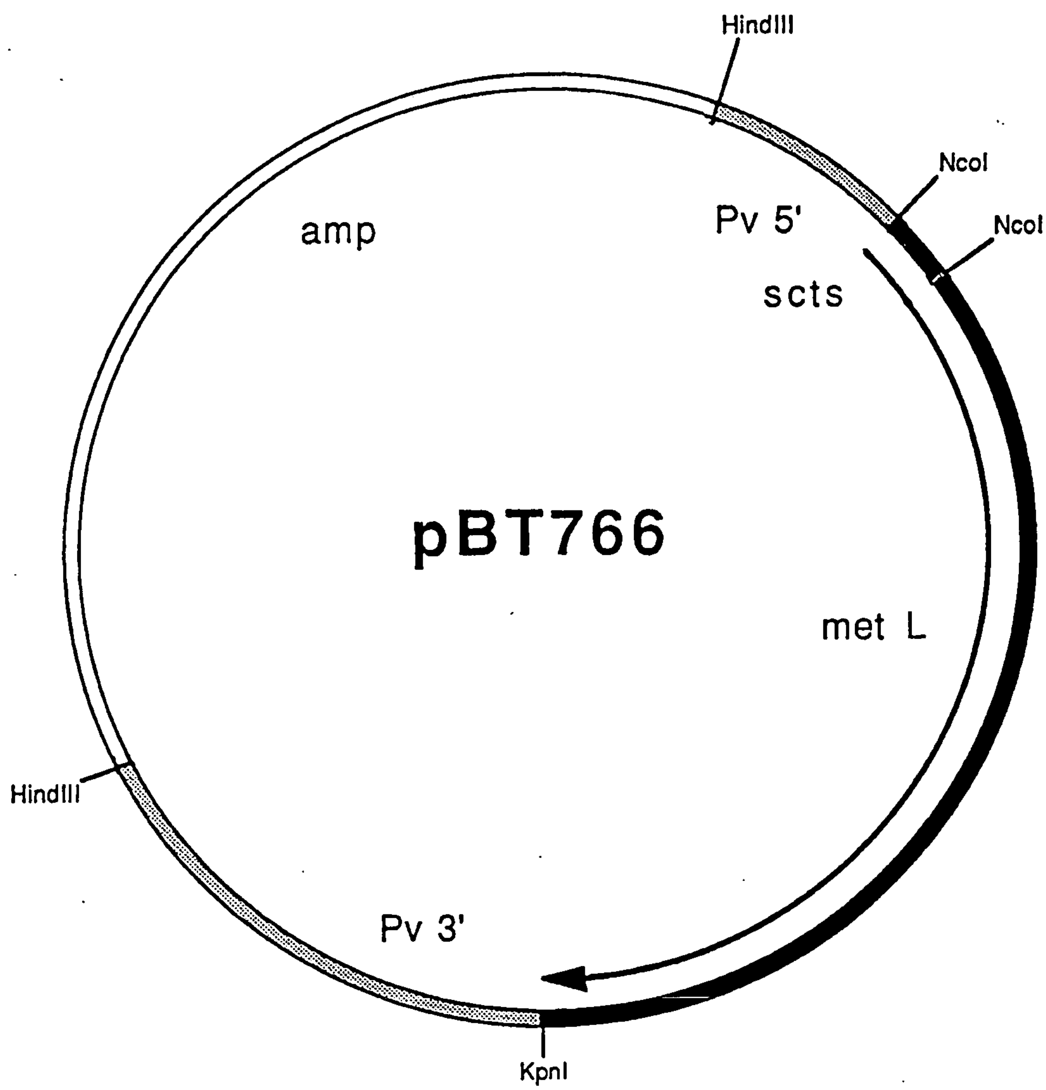


FIG. 6



INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 95/08501

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 C12N15/53 C12N15/54 A01H5/00 A01H5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE WPI Section Ch, Week 9214 Derwent Publications Ltd., London, GB; Class C06, AN 92-113940 & US,A,7 746 705 (MATTHEWS B F ET AL) , 18 February 1992 see abstract ---	1-10
A	PLANT MOLECULAR BIOLOGY, vol. 22, 1993 pages 301-312, WEISEMANN, J.M., ET AL. 'IDENTIFICATION AND EXPRESSION OF A CDNA FROM DAUCUS CAROTA ENCODING BIFUNCTIONAL ASPARTOKINASE-HOMOSERINE DEHYDROGENASE' see the whole document --- -/--	1-10

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
23 November 1995	05. 12. 95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Maddox, A

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PLANT MOLECULAR BIOLOGY, vol. 24, March 1994 pages 835-851, GHISLAIN, M., ET AL. 'MOLECULAR ANALYSIS OF THE ASPARTATE KINASE-HOMOSERINE DEHYDROGENASE GENE FROM ARABIDOPSIS THALIANA' see the whole document	1-10
A	WO,A,93 19190 (DU PONT) 30 September 1993 see page 15 - page 17	1-10
A	THE PLANT JOURNAL, vol. 3, no. 5, 1993 pages 721-727, KARCHI, H., ET AL. 'SEED-SPECIFIC EXPRESSION OF A BACTERIAL DESENSITIZED ASPARTATE KINASE INCREASE THE PRODUCTION OF SEED THREONINE AND METHIONINE IN TRANSGENIC TOBACCO' see the whole document	1-10
A	PLANT MOLECULAR BIOLOGY, vol. 23, 1993 pages 759-768, SHAUL, O., ET AL. 'CONCERTED REGULATION OF LYSINE AND THREONINE SYNTHESIS IN TOBACCO PLANTS EXPRESSING BACTERIAL FEEDBACK-INSENSITIVE ASPARTATE KINASE AND DHIYDRODIPICOLINATE SYNTHASE' see page 765, right column, last paragraph - page 766	1-10

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 95/08501

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9319190	30-09-93	AU-B- 3923393	21-10-93
		CA-A- 2132414	30-09-93
		EP-A- 0640141	01-03-95
		JP-T- 7504821	01-06-95